



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Critical appraisal of CRP measurement for the prediction of coronary heart disease events

**Citation for published version:**

Shah, T, Casas, JP, Cooper, JA, Tzoulaki, I, Sofat, R, McCormack, V, Smeeth, L, Deanfield, JE, Lowe, GD, Rumley, A, Fowkes, FGR, Humphries, SE & Hingorani, AD 2009, 'Critical appraisal of CRP measurement for the prediction of coronary heart disease events: new data and systematic review of 31 prospective cohorts', *International Journal of Epidemiology*, vol. 38, no. 1, pp. 217-31. <https://doi.org/10.1093/ije/dyn217>

**Digital Object Identifier (DOI):**

[10.1093/ije/dyn217](https://doi.org/10.1093/ije/dyn217)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

International Journal of Epidemiology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## CARDIOVASCULAR DISEASE

# Critical appraisal of CRP measurement for the prediction of coronary heart disease events: new data and systematic review of 31 prospective cohorts

Tina Shah,<sup>1†</sup> Juan P Casas,<sup>2†</sup> Jackie A Cooper,<sup>1</sup> Ioanna Tzoulaki,<sup>3</sup> Reecha Sofat,<sup>1</sup> Valerie McCormack,<sup>2</sup> Liam Smeeth,<sup>2</sup> John E Deanfield,<sup>4</sup> Gordon D Lowe,<sup>5</sup> Ann Rumley,<sup>5</sup> F Gerald R Fowkes,<sup>6</sup> Steve E Humphries<sup>1</sup> and Aroon D Hingorani<sup>1,7\*</sup>

---

**Accepted** 3 September 2008

**Background** Non-uniform reporting of relevant relationships and metrics hampers critical appraisal of the clinical utility of C-reactive protein (CRP) measurement for prediction of later coronary events.

**Methods** We evaluated the predictive performance of CRP in the Northwick Park Heart Study (NPHS-II) and the Edinburgh Artery Study (EAS) comparing discrimination by area under the ROC curve (AUC), calibration and reclassification. We set the findings in the context of a systematic review of published studies comparing different available and imputed measures of prediction. Risk estimates per-quantile of CRP were pooled using a random effects model to infer the shape of the CRP-coronary event relationship.

**Results** NPHS-II and EAS (3441 individuals, 309 coronary events): CRP alone provided modest discrimination for coronary heart disease (AUC 0.61 and 0.62 in NPHS-II and EAS, respectively) and only modest improvement in the discrimination of a Framingham-based risk score (FRS) (increment in AUC 0.04 and –0.01, respectively). Risk models based on FRS alone and FRS + CRP were both well calibrated and the net reclassification improvement (NRI) was 8.5% in NPHS-II and 8.8% in EAS with four risk categories, falling to 4.9% and 3.0% for 10-year coronary disease risk threshold of 15%. Systematic review (31 prospective studies 84 063 individuals, 11 252 coronary events): pooled inferred values for the AUC for CRP alone were 0.59 (0.57, 0.61), 0.59 (0.57, 0.61) and 0.57 (0.54, 0.61) for studies of <5, 5–10 and >10 years follow up, respectively. Evidence from 13 studies (7201 cases) indicated that CRP did not consistently improve performance

---

<sup>1</sup> Centre for Clinical Pharmacology, University College London, London, UK.

<sup>2</sup> Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK.

<sup>3</sup> Department of Epidemiology and Public Health, Imperial College London, London, UK.

<sup>4</sup> Vascular Physiology Unit, Institute of Child Health, University College London, UK.

<sup>5</sup> Division of Cardiovascular and Medical Sciences, University of Glasgow, Royal Infirmary, Glasgow, UK.

<sup>6</sup> Wolfson Unit for Prevention of Peripheral Vascular Diseases, Public Health Sciences, University of Edinburgh, Edinburgh, UK.

<sup>7</sup> Department of Epidemiology and Public Health, University College London, London, UK.

<sup>†</sup> These authors contributed equally to this work.

\* Corresponding author. Centre for Clinical Pharmacology, British Heart Foundation Laboratories at UCL, London WC1E 6JJ, UK. E-mail: a.hingorani@ucl.ac.uk

of the Framingham risk score when assessed by discrimination, with AUC increments in the range 0–0.15. Evidence from six studies (2430 cases) showed that CRP provided statistically significant but quantitatively small improvement in calibration of models based on established risk factors in some but not all studies. The wide overlap of CRP values among people who later suffered events and those who did not appeared to be explained by the consistently log-normal distribution of CRP and a graded continuous increment in coronary risk across the whole range of values without a threshold, such that a large proportion of events occurred among the many individuals with near average levels of CRP.

**Conclusions** CRP does not perform better than the Framingham risk equation for discrimination. The improvement in risk stratification or reclassification from addition of CRP to models based on established risk factors is small and inconsistent. Guidance on the clinical use of CRP measurement in the prediction of coronary events may require updating in light of this large comparative analysis.

**Keywords** C-reactive protein, prediction, coronary heart disease, primary prevention, risk stratification

---

## Introduction

Primary prevention of cardiovascular events currently involves targeting interventions to those at high absolute risk, identified using risk-prediction instruments such as the Framingham equation, that integrate information on established risk factors.<sup>1</sup> However, a large proportion of events occur among individuals with near average levels of continuous risk factors,<sup>2</sup> or at intermediate Framingham risk. With emerging evidence on the role of inflammation in atherosclerosis, there is interest in the potential predictive utility of C-reactive protein (CRP), a sensitive circulating biomarker of inflammation.<sup>3</sup> Nearly, 40 reports from prospective studies and three meta-analyses<sup>4–6</sup> indicate a highly consistent, moderate association of CRP with later coronary heart disease (CHD) events among clinically healthy subjects. In 2004, consensus statements from population, laboratory science and clinical practice expert committees convened by the American Heart Association/Centres for Disease Control (AHA/CDC) indicated that 'CRP may be used at the discretion of the physician as part of a global coronary risk assessment in adults without known cardiovascular disease', and that a CRP value above a cut-point of 3 mg/l was indicative of subjects at high risk.<sup>7–10</sup> However, all committees highlighted gaps in the available data and made recommendations for additional research. Although several CRP tests have received FDA approval, recent publications have highlighted ongoing debate on the real utility of CRP measurement in comparison with or as a supplement to conventional risk assessment of

CHD,<sup>11–15</sup> leaving clinicians, policymakers and patients uncertain as to the value of this test.

Reporting of measures of predictive performance in the many prospective studies of CRP has been inconsistent. Although measures of the strength of the association, e.g. using hazard or odds ratios, have achieved greater prominence (and are important in aetiological analysis), these measures provide limited information on predictive utility.<sup>16,17</sup> To complicate matters, there is debate on the most appropriate metric for the evaluation of a new marker. Established measures include discrimination [the ability of a marker to distinguish individuals who will develop an event, assessed by sensitivity and specificity and the area under the ROC curve (AUC)], and calibration (use of a risk model to order or stratify risk, whose accuracy is assessed by tests that evaluate model fit). More recently, concerns that neither of these approaches provides a fair evaluation of a new marker has motivated interest in the development of new measures of predictive utility.<sup>18</sup> A method proposed recently involves quantifying the extent to which a new marker shifts individuals between categories of CHD risk initially determined using established risk models (reclassification).<sup>19</sup>

A 2006 NHLBI workshop on CRP included the following recommendations: (i) 'analyses of individual and pooled data from existing and new epidemiologic studies (or ancillary studies) to examine whether CRP measurement improves CVD risk prediction beyond traditional risk factors and can help target who should be treated'; (ii) research to assess 'whether replacing a traditional risk factor(s) with CRP (or other newly

discovered risk predictors) could improve CVD risk prediction'; and (iii) research on 'the shape of the dose-response curve' (<http://www.nhlbi.nih.gov/meetings/workshops/crp/report.htm>). We therefore assessed the performance of CRP as a risk marker in the Northwick Park Heart Study II (NPHS-II) and replicated the findings in the Edinburgh Artery Study (EAS), two population-based prospective studies. For these analyses we tested discrimination, calibration and reclassification to assess consistency between methods. Next, we conducted a systematic review of prior prospective studies that had examined the association of CRP and CHD to assess consistency between studies. This review differs importantly from prior overviews in this area in that our focus was on the evaluation of the *predictive* performance of CRP, rather than simply its association, and involved inferring measures relevant to prediction where these were unreported. Finally, we examined the population distribution of CRP in all the studies, and the shape of the relationship with CHD events, to better understand the factors that could constrain performance of this marker.

## Methods

### Northwick Park Heart Study II

NPHS-II is a prospective study of 3012 healthy middle-aged men of European descent (age at recruitment 50–64 years), with enrolment commencing in 1989.<sup>20</sup> Nine general practices participated in the study. The study was approved by the local institutional review committee and all subjects provided written informed consent. Further details are provided in the Supplementary Data.<sup>22–24</sup>

### Edinburgh Artery Study

The EAS is a prospective study of 1592 European descent individuals (809 men and 783 women) aged 55–74 years enrolled in 1988. Individuals were selected at random from 11 general practices serving a range of socio-economic and geographical areas throughout the city of Edinburgh, Scotland. Details of the study recruitment and examination process have been described elsewhere.<sup>21</sup> Further details are provided in the Supplementary Data.<sup>25</sup>

### Statistical analyses in NPHS-II and EAS

#### Association

Log-transformations were conducted for non-normally distributed data and for these variables means are geometric and standard deviations approximate. The association of CRP and risk of CHD events is reported as a hazard ratio with 95% CI obtained from a Cox proportional hazards model. Age-adjusted incidence rates (cases/1000 person years) and the absolute number of events are reported by tertile of CRP and in categories defined by CRP cut-points <1, 1–3 and >3 mg/l. Means and standard

deviations (SD) of log-transformed CRP in those who suffered CHD events and those who remained event-free were used to construct normal curves to determine the overlap between the distributions in the two groups.

#### Model discrimination

The performance of CRP for discrimination was assessed by means of the disease detection rate (DR) or sensitivity, i.e. the proportion of those who developed events who tested positive using CRP cut-points corresponding to pre-set false-positive rates of 5% (DR5) or 10% (DR10). In addition, we calculated the positive predictive value of a CRP cut-point of 3 mg/l i.e. the proportion of people with a CRP value >3 mg/l who suffered an event. The incremental effect on discrimination of adding CRP to the Framingham risk score was summarized by means of the AUC and Harrell's *C*-statistic. The Harrell's *C*-statistic from a Cox model is conceptually analogous to the AUC estimated from logistic models which allows for right-censored data and variable time to follow up.<sup>16,26–28</sup>

#### Model calibration

The performance of risk estimates generated from Cox models using Framingham variables alone,<sup>29</sup> or with the addition of CRP, was evaluated by ranking participants according to one-fifths of predicted risk, and comparing within each category, the predicted and observed 10-year risks of CHD in NPHS-II and 17-year risks of CHD in EAS. The Akaike's information criterion (AIC) and Bayes information criterion (BIC) were used to assess global fit using Framingham variables alone or with CRP. The Hosmer–Lemeshow chi-square statistic<sup>30</sup> was used to assess model fit. Small chi-square values indicate a good calibration while values exceeding 20 ( $P < 0.01$ ) indicate significant lack of calibration.

#### Reclassification of CHD risk

The extent to which CRP reassigned individuals to risk categories that better reflected their final outcome was assessed using the NRI measure.<sup>19</sup> For individuals who develop an event, risk classification is considered improved if the individual moves to a higher risk category with the addition of a marker, and worsened if the individual moves to a lower one. For individuals who remain healthy the converse is true. For people who develop events, the difference in the proportion of individuals moving up and the proportion moving down a category is calculated. For people remaining healthy, the proportion of individuals moving down minus the proportion moving up a category is calculated. The NRI is obtained as the sum of these two values. First, individuals were assigned to one of the following 10-year CHD risk categories based on a model using established risk factors: 0% to <5%, 5% to <10%, 10% to <15% and ≥15%, similar to previous reports.<sup>31</sup> Next, the NRI



was calculated following the addition of CRP to the risk model. In a separate analysis, individuals were assigned to one of two risk categories based on a 10-year CHD risk of 0% to <15% or  $\geq 15\%$ . This allowed us to assess the potential effect of reclassification on the clinical decision to prescribe cholesterol lowering therapy for primary prevention based on thresholds for intervention advocated by UK guidelines (e.g. NICE guidelines on primary prevention with statins, <http://www.nice.org.uk/guidance/>). In addition, to facilitate comparison with two prior reports, reclassification tables were drawn to allow comparison of the observed (actual) risk in each category of predicted risk comparing models that differed only in their inclusion or omission of CRP.

## Systematic review

### Search strategy and data extraction

Two electronic databases (Medline and EMBASE) were searched up to and including August 2007 for all prospective studies (including cohort, nested case-control or case-cohort studies) of initially healthy subjects evaluating the association between CRP concentration and coronary events with no threshold sample size. For the search, the MeSH terms, 'C-reactive protein' and 'CRP' in combination with 'coronary', 'coronary heart disease', 'CHD' and 'CVD' were used and the search was limited by the terms 'human' and 'English Language'. Studies in which total mortality was the only outcome reported were excluded. See Supplementary Data and Supplementary Figure 1 for further details.

### Analysis

Measures relevant to the performance of a screening test tended not to be reported in many of the previously reported studies and we estimated some of them as follows. We first extracted values of the geometric mean CRP (and approximate SD) separately for incident cases and those remaining free of events (controls). Where the median was reported, this was used as an estimate of the geometric mean and approximate SD was calculated from the inter-quartile range, assuming a log-normal distribution. Where no measure of dispersion was reported, an estimated SD based on pooled values from studies that provided this measure was applied. These data were then used to construct the log-normal distributions of CRP from each study for the cases and controls separately, which permitted estimation of study specific values for the DR (sensitivity) for a 5% or 10% false-positive rate (DR5 and DR10), as well as the AUC for CRP using methods reported previously.<sup>17,27</sup> Where appropriate, study-specific values for AUC and DR were pooled using a random-effects model with weighting by the inverse of the variance.

## Results

### NPHS-II and EAS

#### Association of CRP values with incident CHD

There were 162 incident cases of CHD in NPHS-II and 147 in EAS among the individuals with valid CRP measurements. Higher CRP concentrations were associated with a higher risk of CHD events in both NPHS-II and EAS. In NPHS-II a hazard ratio of 2.22 (95% CI 1.50, 3.30) and EAS a hazard ratio of 1.87 (95% CI 1.21, 2.89) was observed for the comparison of the top vs bottom tertile of the CRP distribution in age- and primary-care recruitment centre-adjusted models. These values were similar in magnitude to those reported in a previous meta-analysis.<sup>6</sup>

#### Effect of CRP on discrimination of CHD

With a CRP cut-point corresponding to a pre-set false-positive rate of 5% (DR5), the sensitivity for detection of CHD events was 11.2% in NPHS-II and 10.9% in EAS, i.e. around 10% of individuals who suffered CHD events were identified. With a CRP cut-point corresponding to a 10% false-positive rate, the DR (DR10) was 17.8% in NPHS-II and 19.7% for EAS. Using the cut-point of 3 mg/l, the positive predictive value was 8.7% in NPHS-II and 20.0% in EAS. The AUC for CRP in NPHS-II was 0.61 (95% CI 0.57, 0.66), and in EAS was 0.62 (95% CI 0.57, 0.67). These estimates of the AUC were modified only minimally when account was made of the effect of time on the accumulation of CHD events using Harrell's C-statistic (Table 1). Values of the C-statistic for the Framingham-based model alone and the Framingham plus CRP model were 0.62 (0.60, 0.65) and 0.66 (0.63, 0.68) in NPHS-II, and 0.68 (0.64, 0.71) and 0.67 (0.63, 0.71) in EAS, respectively (Table 1).

#### Effect of CRP on calibration of risk models based on Framingham variables

Two risk models were compared in NPHS-II and EAS: a Cox model derived using Framingham variables, and the same model with the addition of CRP. Individuals were first divided into one-fifths of predicted risk and observed event rates were compared with those expected. The global fit of the model was improved by the addition of CRP when assessed by the AIC and the BIC, which in NPHS and EAS were 2368.6 and 2397.6 for Framingham variables, respectively, and 2355.8 and 2390.5 for Framingham variables plus CRP, respectively ( $P < 0.001$ ). However, because statistically significant improvements in model fit do not necessarily correspond to large absolute differences, we calculated the ratio of predicted to observed event-rates in each one-fifth of risk in the two models generated from both studies (Figure 1). This indicated that the magnitude of improvement in calibration conferred by adding CRP to the Framingham model was small and this was consistent with the Hosmer-Lemeshow  $P$ -value which

**Table 1** Relationship between hazard ratio for CHD, AUC and the C-statistic for CRP, established risk predictors, the Framingham risk score and the Framingham risk score and CRP combined in NPHS II and EAS

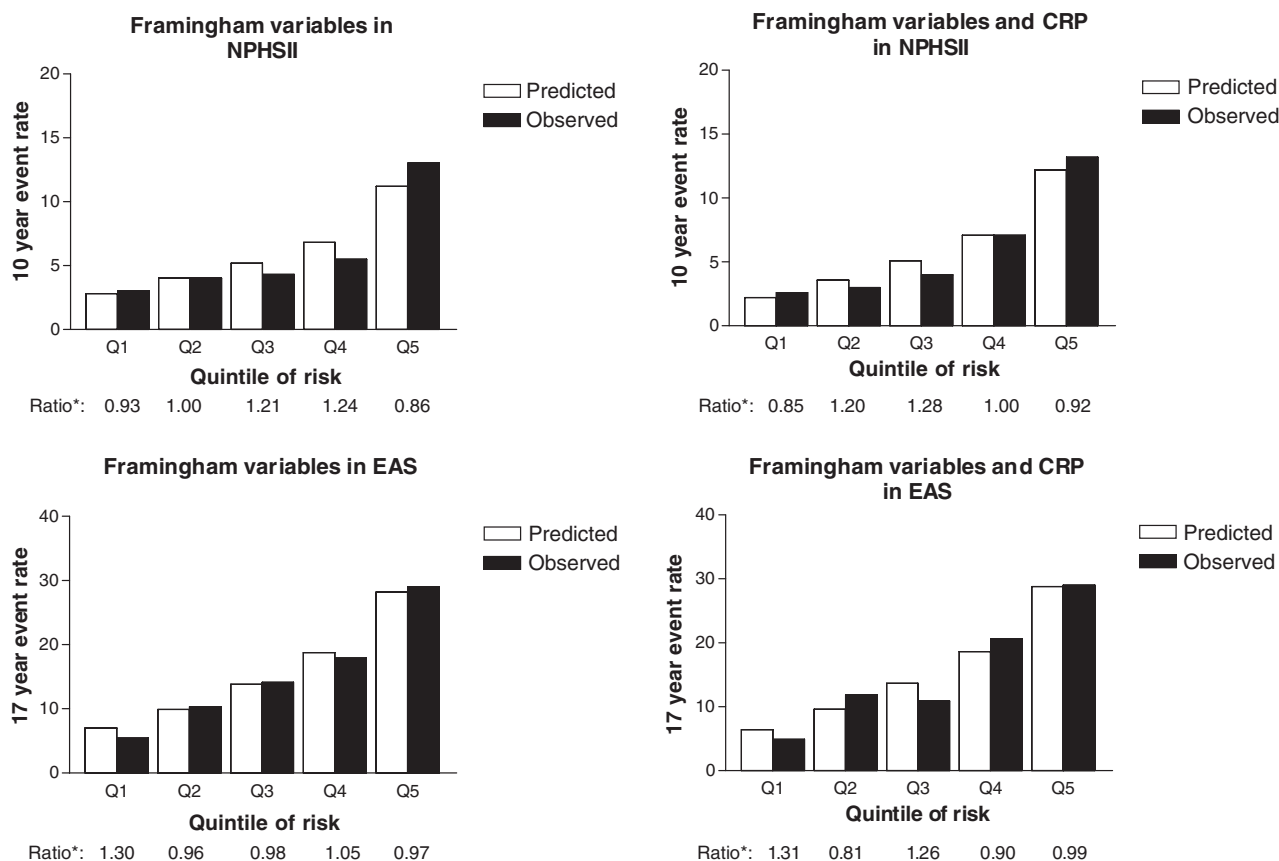
|         | NPHS-II                              |                     |                                 | EAS                                  |                     |                                 |
|---------|--------------------------------------|---------------------|---------------------------------|--------------------------------------|---------------------|---------------------------------|
|         | Hazard ratio <sup>a</sup><br>(95%CI) | AUC<br>(95%CI)      | C-index <sup>b</sup><br>(95%CI) | Hazard ratio <sup>a</sup><br>(95%CI) | AUC<br>(95%CI)      | C index <sup>b</sup><br>(95%CI) |
| CRP     | 2.53<br>(1.44–4.62)                  | 0.61<br>(0.57–0.66) | 0.63<br>(0.60–0.66)             | 1.87<br>(1.21–2.89)                  | 0.62<br>(0.57–0.67) | 0.59<br>(0.54–0.63)             |
| SBP     | 1.96<br>(1.10–3.50)                  | 0.60<br>(0.55–0.64) | 0.59<br>(0.56–0.62)             | 2.00<br>(1.39–2.89)                  | 0.62<br>(0.58–0.66) | 0.58<br>(0.54–0.62)             |
| TC      | 1.35<br>(0.75–2.43)                  | 0.59<br>(0.55–0.64) | 0.59<br>(0.56–0.62)             | 1.91<br>(1.32–2.76)                  | 0.62<br>(0.58–0.66) | 0.55<br>(0.51–0.59)             |
| HDL-C   | 0.74<br>(0.55–0.99)                  | 0.57<br>(0.50–0.65) | 0.58<br>(0.53–0.62)             | 0.56<br>(0.37–0.84)                  | 0.61<br>(0.57–0.65) | 0.61<br>(0.57–0.64)             |
| FRS     | 1.47<br>(0.68–3.18)                  | 0.62<br>(0.57–0.66) | 0.62<br>(0.60–0.65)             | 4.61<br>(2.98–7.13)                  | 0.65<br>(0.61–0.70) | 0.68<br>(0.64–0.71)             |
| FRS+CRP | 2.11<br>(1.04–4.31)                  | 0.64<br>(0.60–0.69) | 0.66<br>(0.63–0.68)             | 3.67<br>(2.16–6.23)                  | 0.65<br>(0.61–0.70) | 0.67<br>(0.63–0.71)             |

Framingham variables used in NPHS-II were age, smoking, SBP, cholesterol, diabetes.

Framingham variables used in EAS were age, sex, smoking, SBP, cholesterol, HDL, diabetes.

<sup>a</sup>Hazard ratio for top vs bottom tertile adjusted for age, practice, blood pressure, total- and HDL- cholesterol where appropriate.

<sup>b</sup>Predictive accuracy is calculated as Harrell's C-index, which extends the AUC to the case of right-censored survival data.



\*Ratio of predicted/observed event rate.

**Figure 1** Predicted and observed risk of CHD based on Cox models derived using traditional risk factors alone and with the addition of CRP

indicated that in NPHS-II and EAS studies, models using Framingham variables alone and Framingham variables plus CRP were both well calibrated. In NPHS-II study the Hosmer–Lemeshow *P*-values were 0.82 for Framingham without CRP and 0.90 including CRP, and in the EAS cohort, the *P*-values were 0.65 for Framingham without CRP and 0.65 including CRP.

#### ***Effect of CRP on prediction using the NRI measure***

The reclassification of individuals following the addition of CRP to a Framingham-based model is summarized in Table 2 according to the eventual outcome. In NPHS-II, after generation of four 10-year CHD risk categories, a total of 716 men (60 cases and 656 controls) were reclassified, of which 368 (51.4%) moved to higher categories and 348 (48.6%) to lower categories. Of the 60 cases reclassified, 23 (38%) were inappropriately reclassified to *lower* risk categories and a similar proportion of inappropriate reclassifications was observed among controls, with 331 (50.5%) out of 656 controls being reclassified to *higher* categories. The NRI was 8.5% (−1.3, 18.3; *P* = 0.09). Very similar results were observed in the EAS study when using the same four risk categories (Table 2, panel A). When two risk-categories were generated using a treatment threshold for primary prevention of 15% CHD risk at 10-years, the number of subjects correctly reclassified by case–control status was substantially lower. In NPHS-II, in only 10 cases and 18 controls would the addition of CRP have impacted on the decision to prescribe statins, with a similarly low number in the EAS study. Consequently, the NRI using a treatment threshold of 15% was reduced to 4.9% (0.8–9.0) in NPHS-II and 3.0% (−3.0 to 9.2) in EAS (Table 2, Panel B).

#### ***Distribution of CRP values among cases and controls***

In both EAS and NPHS-II, the distribution of CRP was right-skewed. Log-transformation resulted in normalization of the distribution in both studies, with the geometric mean being 2.46 mg/l (approximate SD 2.50 mg/l) in NPHS-II, and 1.93 mg/l (approximate SD 3.02 mg/l) in EAS. In NPHS-II, the geometric mean CRP (approximate SD) among the 162 men who later developed CHD was 3.66 (3.66) mg/l, and 2.40 (2.44) mg/l among the 2317 men who remained event-free. In EAS, geometric mean CRP (approximate SD) among the 147 individuals who later developed CHD was 2.47 (3.01) mg/l, and was 1.85 (3.00) mg/l among the 815 people who remained event free. Though on average CRP was higher among individuals who later suffered events, there was substantial overlap of the log-normal distributions of CRP values between the two groups (Supplementary Figure 2), imposing a difficulty in setting a threshold value for CRP that distinguishes individuals who later suffer events. Although the *risk* of CHD events increased with higher CRP values, 52.4% of CHD events in NPHS-II and 57.1% events in EAS occurred

among individuals in the middle and lower tertiles of the CRP distribution (Supplementary Table 1). When the CRP distribution was divided according to the cut-points <1, 1–3 and >3 mg/l, recommended as defining individuals at low, intermediate and high risk, respectively, a similar percentage of events (41.3% in NPHS-II and 57.1% in EAS) were observed among individuals at low or intermediate risk categories.

### **Systematic review**

#### ***Reporting of metrics of prediction performance in published prospective studies***

Thirty-one studies of 28 prospective cohorts involving a total of 84 063 individuals and 11 252 incident CHD events were identified up to August 2007 (Supplementary Table 4; Supplementary references W1–W31). Twelve (38.7%) were studies of men alone (Supplementary references W1, W3, W4, W9–W13, W17, W18, W21, W25), six (19.4%) were studies of women alone (Supplementary references W5, W6, W16, W21, W23, W27) and 13 (41.9%) studies of both men and women (Supplementary references W2, W7, W8, W14, W15, W19, W20, W22, W24, W28, W29, W31). All 31 studies provided a measure of strength of association. Two studies (6.7%; 20 191 individuals, 531 incident cases) reported the AUC for CRP alone (Supplementary references W6, W24). A total of 13 studies from the 31 prospective cohorts (42%; 84 292 individuals, 7201 incident cases) reported the effect of adding CRP to the Framingham model on discrimination (Supplementary references W6, W8, W13, W15, W19, W20, W22–W25, W27, W28, W30). One study (15 048 individuals, 390 incident cases) assessed reclassification but did not report the NRI measure (Supplementary references W27). Up to 2003, only one study had reported on metrics related directly to predictive performance (Supplementary reference W6).

#### ***Shape of the CHD association and absolute number of events per CRP quantile in published prospective studies***

Thirty studies (96.8%) reported on the shape of the CRP distribution, and in all these studies it was reported to be skewed (Supplementary references W1–W3, W5–W31). Twenty-one studies (67.7%) reported that logarithmic transformation normalized the distribution, (Supplementary references W1–W3, W7–W15, W17, W20, W22–W24, W28–W31) as observed in NPHS-II and EAS. Including NPHS-II and EAS, a total of 12 studies reporting risk according to quartiles of CRP concentration, and 10 studies reporting risk according to quintiles, allowed estimation of the shape of the relationship between log-CRP and log risk of CHD (Supplementary references W1, W3, W5, W6, W10, W12–W14, W16–W19, W21, W22, W26). There were no systematic differences in the characteristics of studies reporting CRP effects by quartiles or quintiles

**Table 2** NRI in the NPHS-II and EAS studies using Framingham-based models with and without CRP: (A) using four 10-year CHD risk categories and (B) using categories based on a threshold for intervention of a 10-year CHD risk of >15%**Panel A****Predicted 10-year CHD risk****Framingham variables plus CRP**

|  | No. of individuals |       |        |      | Reclassified   |                | Net correctly reclassified |
|--|--------------------|-------|--------|------|----------------|----------------|----------------------------|
| Framingham variables   | <5%                | 5–10% | 10–15% | >15% | Increased risk | Decreased risk |                            |
| NPHS-II study  |                    |       |        |      |                |                |                            |
| Individuals without CHD events during follow-up ( <i>n</i> = 2253) |                    |       |        |      |                |                |                            |
| 0–<5%  | 890                | 181   | 0      | 0    | 331            | 325            | –0.3%                      |
| 5–10%  | 242                | 604   | 115    | 4    |                |                |                            |
| 10–15%   | 2                  | 63    | 73     | 31   |                |                |                            |
| ≥ 15%  | 0                  | 2     | 16     | 30   |                |                |                            |
| Individuals with CHD events during follow-up ( <i>n</i> = 159)     |                    |       |        |      |                |                |                            |
| 0–<5%  | 38                 | 10    | 0      | 0    | 37             | 23             | 8.8%                       |
| 5–10%  | 12                 | 40    | 17     | 2    |                |                |                            |
| 10–15%   | 0                  | 10    | 12     | 8    |                |                |                            |
| ≥ 15%  | 0                  | 0     | 1      | 9    |                |                |                            |
| NRI (95% CI)   |                    |       |        |      |                |                | 8.5% (–1.3, 18.3)          |
| Edinburgh Artery Study   |                    |       |        |      |                |                |                            |
| Individuals without CHD events during follow-up ( <i>n</i> = 777)  |                    |       |        |      |                |                |                            |
| 0–<5%  | 247                | 23    | 0      | 0    | 83             | 124            | 5.3%                       |
| 5–10%  | 59                 | 219   | 32     | 2    |                |                |                            |
| 10–15%   | 0                  | 46    | 56     | 26   |                |                |                            |
| ≥ 15%  | 0                  | 2     | 17     | 48   |                |                |                            |
| Individuals with CHD events during follow-up ( <i>n</i> = 142)     |                    |       |        |      |                |                |                            |
| 0–<5%  | 18                 | 5     | 0      | 0    | 26             | 21             | 3.5%                       |
| 5–10%  | 9                  | 29    | 9      | 0    |                |                |                            |
| 10–15%   | 0                  | 6     | 27     | 12   |                |                |                            |
| ≥ 15%  | 0                  | 1     | 5      | 21   |                |                |                            |
| NRI (95%CI)  |                    |       |        |      |                |                | 8.8% (–1.3 to 18.9)        |

**Panel B****Predicted 10-year CHD risk****Framingham variables plus CRP**

| Framingham variables plus CRI                                      | No. of individuals |       | Reclassified   |                | Net correctly reclassified |
|--|--------------------|-------|----------------|----------------|----------------------------|
| Framingham variables   | 0-15%              | ≥ 15% | Increased risk | Decreased risk |                            |
| <b>NPHS-II study</b>   |                    |       |                |                |                            |
| Individuals without CHD events during follow-up ( <i>n</i> = 2253) |                    |       |                |                |                            |
| 0-15%  | 2170               | 35    | 35             | 18             | -0.8%                      |
| ≥ 15%  | 18                 | 30    |                |                |                            |
| Individuals with CHD events during follow-up ( <i>n</i> = 159)     |                    |       |                |                |                            |
| 0-15%  | 139                | 10    | 10             | 1              | 5.7%                       |
| ≥ 15%  | 1                  | 9     |                |                |                            |
| NRI (95%CI)  |                    |       |                |                | 4.9% (0.8, 9.0)            |
| <b>Edinburgh Artery Study</b>                                      |                    |       |                |                |                            |
| Individuals without CHD events during follow-up ( <i>n</i> = 777)  |                    |       |                |                |                            |
| 0-15%  | 682                | 28    | 28             | 19             | -1.2%                      |
| ≥ 15%  | 19                 | 48    |                |                |                            |
| Individuals with CHD events during follow-up ( <i>n</i> = 142)     |                    |       |                |                |                            |
| 0-15%  | 103                | 12    | 12             | 6              | 4.2%                       |
| ≥ 15%  | 6                  | 21    |                |                |                            |
| NRI (95%CI)  |                    |       |                |                | 3.0% (-3.0, 9.2)           |



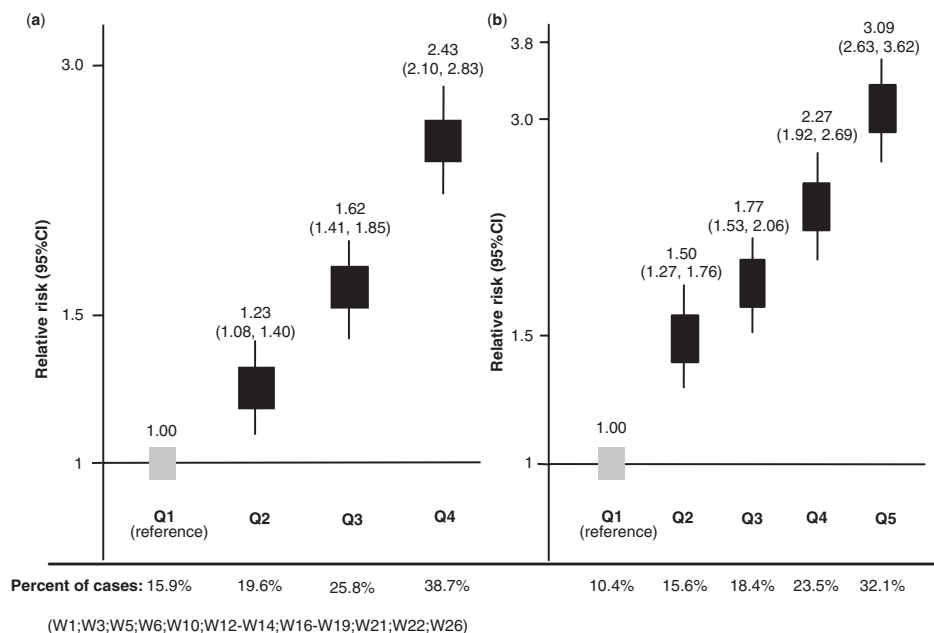
(Supplementary Table 2). A linear graded association of log-CRP values with log risk of events was noted in all studies (Figure 2). The log-normal distribution of CRP in populations, coupled with the graded incremental association of CRP concentration with risk over the whole range of CRP values, leads to the expectation of a substantial proportion of all CVD events occurring among the large number of individuals with near average levels of CRP. In the eight studies, including NPHS-II and EAS, that reported the absolute number of events by tertile; 1772 of 3152 events (56%) occurred among individuals in the lower and middle tertiles of the CRP distribution (Supplementary references W2, W8, W9, W21, W22, W28). In the seven studies, including NPHS-II and EAS (Supplementary references W1, W12, W13, W17, W22), that reported absolute events by quartile, 1105 of 2428 events (46%) occurred among people in the middle two quartiles. These proportions were almost identical to those estimated from the pooled relative risks by quartiles or quintiles from 18 studies with data available (12 412 individuals, 4651 incident cases) (Figure 2).

#### *Derivation and reporting of measures of predictive performance of CRP in published prospective studies*

Ten studies reported on geometric mean CRP and its approximate SD separately among incident cases and those remaining event-free (controls), and seven studies reported separately on median CRP values and the inter-quartile range from which the geometric mean CRP and approximate SD were estimated under the assumption of a log-normal distribution. For the remainder of studies where a measure of dispersion

was not reported, a pooled SD was applied. These were then used to reconstruct the log-normal distribution of baseline CRP values in the two groups defined by eventual outcome. In all the 25 studies with relevant information (40 684 individuals and 9351 cases), there was a substantial overlap of baseline CRP values among incident cases and controls (Supplementary Figure 2), similar to that observed in NPHS-II and EAS.

Estimated disease DRs for a 5% false-positive rate inferred from these distributions are tabulated (Table 3) and those for a 10% false-positive rate shown graphically (Figure 3a), together with estimates of the AUC for CRP from individual studies (Figure 3b). The estimates were consistent with the corresponding values from NPHS-II and EAS. Because it was not possible to estimate the C-statistic from the summary data, we examined the influence of length of follow up on the AUC. There was a broad similarity in the DR10% or AUC estimates in analysis stratified by mean duration of follow up. The seven studies that reported on the geometric mean but provided no measure of dispersion and the single study that reported the median but provided no inter-quartile range were similar in other respects to those that reported these data (Supplementary Table 3). In the 13 studies that reported on the effect on the ROC curve or C-statistic of adding CRP to the Framingham-based models (Supplementary references W6, W8, W13, W15, W19, W20, W22–W25, W27, W28, W30) five reported no change and eight reported an improvement in the AUC ranging from 0.01 to 0.15 (Figure 4).



**Figure 2** Pooled relative risk of CHD: (a) by quartiles and (b) by quintiles of CRP concentration from 12 and 10 prospective studies, respectively. The pooled values were calculated using the inverse variance-weighted method under random effect models

Five studies (16.1%; 42 141 individuals, 2430 incident cases) reported the effect of adding CRP to the *calibration* of risk models utilizing traditional risk factors (Supplementary Table 4; Supplementary references W15, W23, W27–W29). The effect of adding CRP to a base model was assessed in a variety of ways across studies, precluding pooled analyses. However,

regardless of the metric used or reported statistical significance, absolute improvements if present were small.

Two reports from the Women's Health Study (Supplementary reference W27, 38), examined the effect of adding CRP to models based on established risk factors using reclassification tables. Both studies

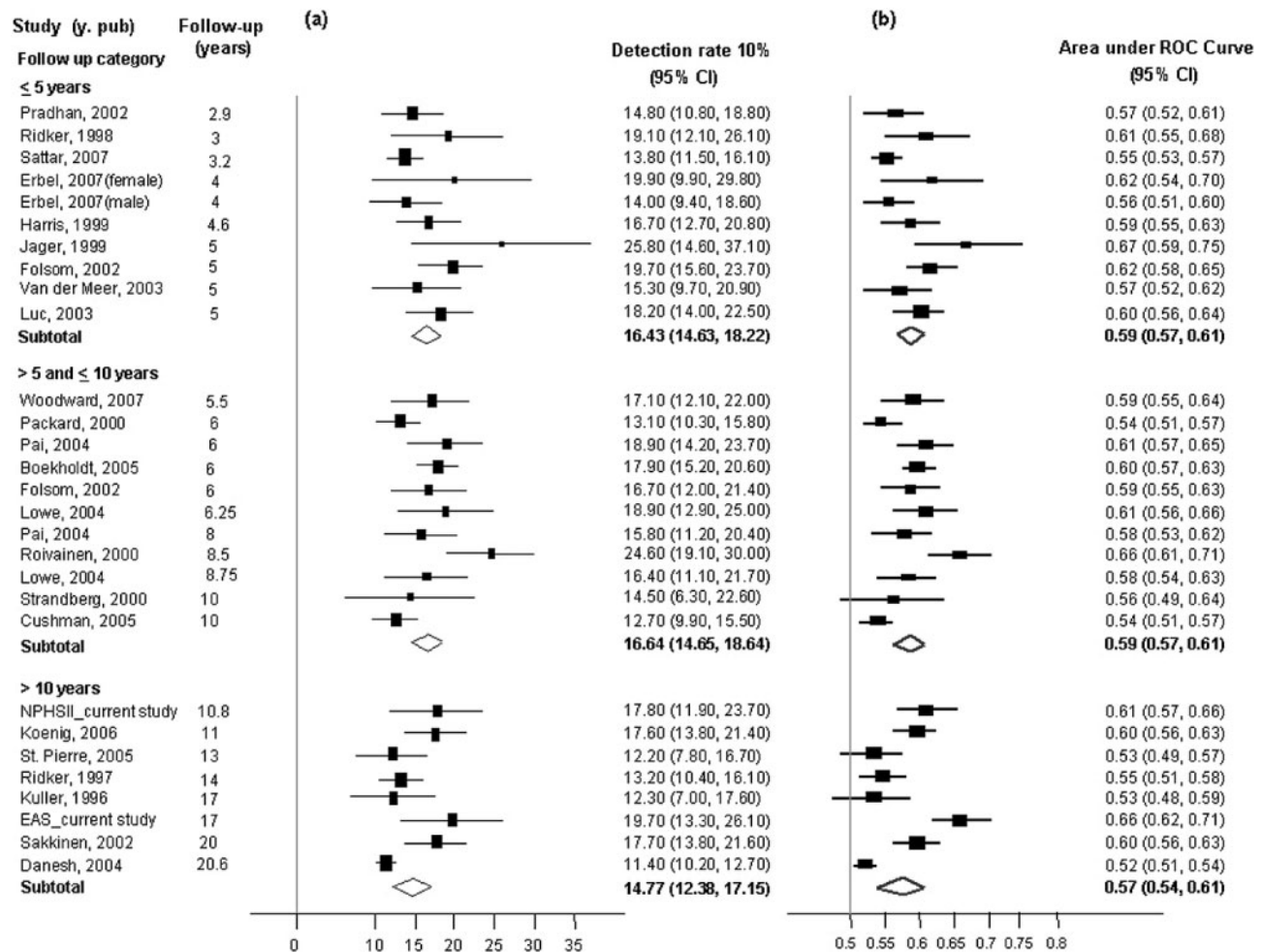
**Table 3** Prospective studies from which measures of discrimination were inferred

| Study<br>(year pub)    | Data set                       | Men<br>(%) | Mean<br>follow-<br>up (yrs) | Primary<br>endpoint | Geometric mean<br>CRP cases <sup>a</sup><br>(mg/l) (SD) | Number<br>cases | Geometric mean<br>CRP controls <sup>a</sup><br>(mg/l) (SD) | Number<br>controls | Mean DR <sub>s</sub><br>95%CI <sup>b</sup><br>(computed) |
|------------------------|--------------------------------|------------|-----------------------------|---------------------|---|-----------------|--|--------------------|--|
| Kuller (1996)          | MRFIT                          | 100        | 17                          | CHD                 | 3.4 <sup>b</sup>  | 148             | 2.9 <sup>b</sup>   | 296                | 6.3 (2.4–10.3)   |
| Ridker (1997)          | PHS                            | 100        | 14                          | CVD                 | 1.37 <sup>b</sup>                                       | 543             | 1.10 <sup>b</sup>  | 543                | 6.9 (4.8–9.1)  |
| Ridker (1998)          | WHS                            | 0          | 3                           | CVD                 | 6.45 <sup>a</sup>                                       | 122             | 3.75 <sup>a</sup>  | 244                | 10.8 (5.3–16.3)  |
| Jager (1999)           | Hoorn                          | 48         | 5                           | CVD<br>mortality    | 3.33 <sup>a</sup> (4.1)                                 | 58              | 1.68 <sup>a</sup> (1.9)                                    | 573                | 15.5 (6.2–24.9)  |
| Packard (2000)         | WOSCOPS                        | 100        | 6                           | CHD                 | 2.36 (2.81)   | 580             | 1.88 (2.92)  | 1160               | 6.8 (4.8–8.9)  |
| Roivainen<br>(2000)    | HHS                            | 100        | 8.5                         | CHD                 | 4.44 <sup>b</sup>                                       | 241             | 2.01 <sup>b</sup>  | 241                | 14.6 (10.2–19.1)   |
| Pradhan (2002)         | WHI-OS                         | 0          | 2.9                         | CHD                 | 3.3 <sup>a</sup> (4.2)                                  | 304             | 2.5 <sup>a</sup> (2.7)                                     | 304                | 7.9 (4.9–11.0)   |
| Sakkinen (2002)        | Honolulu                       | 100        | 20                          | CHD                 | 0.81 (0.7)  | 369             | 0.60 (0.5)   | 1348               | 9.8 (6.8–12.9)   |
| Folsom (2002)          | ARIC<br>(87–93)                | 66         | 6                           | CHD                 | 3.5 <sup>b</sup>  | 242             | 2.3 <sup>b</sup>   | 590                | 9.2 (5.5–12.8)   |
| Folsom (2002)          | ARIC<br>(90–95)                | 66         | 5                           | CHD                 | 2.3 <sup>b</sup>  | 373             | 1.3 <sup>b</sup>   | 590                | 11.2 (8.0–14.3)  |
| Luc (2003)             | PRIME                          | 100        | 5                           | CHD                 | 2.0 <sup>a</sup> (2.1)                                  | 317             | 1.33 <sup>a</sup> (1.5)                                    | 609                | 10.2 (6.8–13.5)  |
| Van der Meer<br>(2003) | Rotterdam                      | 45.5       | 5                           | CHD                 | 2.18 <sup>a</sup> (3.0)                                 | 157             | 1.68 <sup>a</sup> (1.6)                                    | 500                | 8.2 (3.9–12.5)   |
| Lowe (2004)            | Caerphilly                     | 100        | 8.75                        | CHD                 | 2.26 (3.5)  | 189             | 1.66 (1.8)   | 1334               | 9 (4.9–13.0)   |
| Lowe (2004)            | Speedwell                      | 100        | 6.25                        | CHD                 | 2.36 (2.9)  | 162             | 1.5 (1.9)  | 1528               | 10.7 (5.9–15.4)  |
| Danesh (2004)          | Reykjavik                      | 70.3       | 20.6                        | CHD                 | 1.75 (5.3)  | 2459            | 1.28 (5.2)   | 3969               | 5.8 (4.9–6.8)  |
| Pai (2004)             | NHS                            | 0          | 8                           | CHD                 | 3.1 <sup>a</sup> (4.6)                                  | 239             | 2.2 <sup>a</sup> (3.0)                                     | 469                | 8.6 (5.0–12.1)   |
| Pai (2004)             | HPFS                           | 100        | 6                           | CHD                 | 1.68 <sup>a</sup> (1.8)                                 | 265             | 1.08 <sup>a</sup> (1.4)                                    | 529                | 10.7 (6.9–14.4)  |
| Cushman (2005)         | CHS                            | 42         | 10                          | CHD                 | 1.98 <sup>b</sup>                                       | 547             | 1.64 <sup>b</sup>  | 3424               | 6.6 (4.5–8.7)  |
| St. Pierre (2005)      | QCS                            | 100        | 13                          | CHD                 | 2.2 (3.2)   | 210             | 1.8 (3.1)  | 1772               | 6.3 (3.0–9.6)  |
| Boekholdt (2005)       | EPIC-N                         | 63         | 6                           | CHD                 | 2.1 <sup>a</sup> (2.5)                                  | 770             | 1.4 <sup>a</sup> (1.8)                                     | 1515               | 9.9 (7.8–12.1)   |
| Koenig (2006)          | MONICA                         | 55         | 11                          | CHD                 | 2.6 (2.4)   | 382             | 1.4 (2.65)   | 1980               | 9.8 (6.8–12.7)   |
| Sattar (2007)          | PROSPER                        | 48         | 3.2                         | CVD                 | 3.64 (3.08)   | 865             | 3.01 (3.05)  | 4815               | 7.3 (5.6–9.0)  |
| Erbel (2007)           | Heinz<br>Nixdorf               | 100        | 4                           | CAD                 | 1.7 <sup>a</sup>  | 218             | 1.3 <sup>a</sup>   | 1918               | 7.4 (3.9–10.9)   |
| Erbel (2007)           | Heinz<br>Nixdorf               | 0          | 4                           | CAD                 | 2.5 <sup>a</sup>  | 62              | 1.4 <sup>a</sup>   | 2148               | 11.3 (3.4–19.2)  |
| Woodward<br>(2007)     | Fletcher<br>Challenge<br>study | 72         | 5.5                         | CHD                 | 3.81 (9.9)  | 220             | 2.01 (2.98)  | 411                | 9.4 (5.5–13.3)   |

<sup>a</sup>Median CRP values and approximate SD calculated from interquartile range. <sup>b</sup>SD not available.

MRFIT, Multiple Risk Factor Intervention Trial; PHS, Physicians Health Study; WHS, Womens Health Study; WOSCOPS, West of Scotland Coronary Prevention Study; HHS, Helsinki Heart Study; WHI-OS, Womens Health Initiative Observational Study; PRIME, Prospective Epidemiological Study of Myocardial Infarction; NHS, Nurses Health Study; HPFS, Health Professionals Follow-up Study; CHS, Cardiovascular Health Study; QCS, Quebec Cardiovascular Study; EPIC-N, European Prospective Investigation of Cancer at Norfolk; MONICA, MONitoring in CArdiovascular diseases; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk.

(Supplementary references W1–W3;W5–W14;W16–W22;W26–W31)



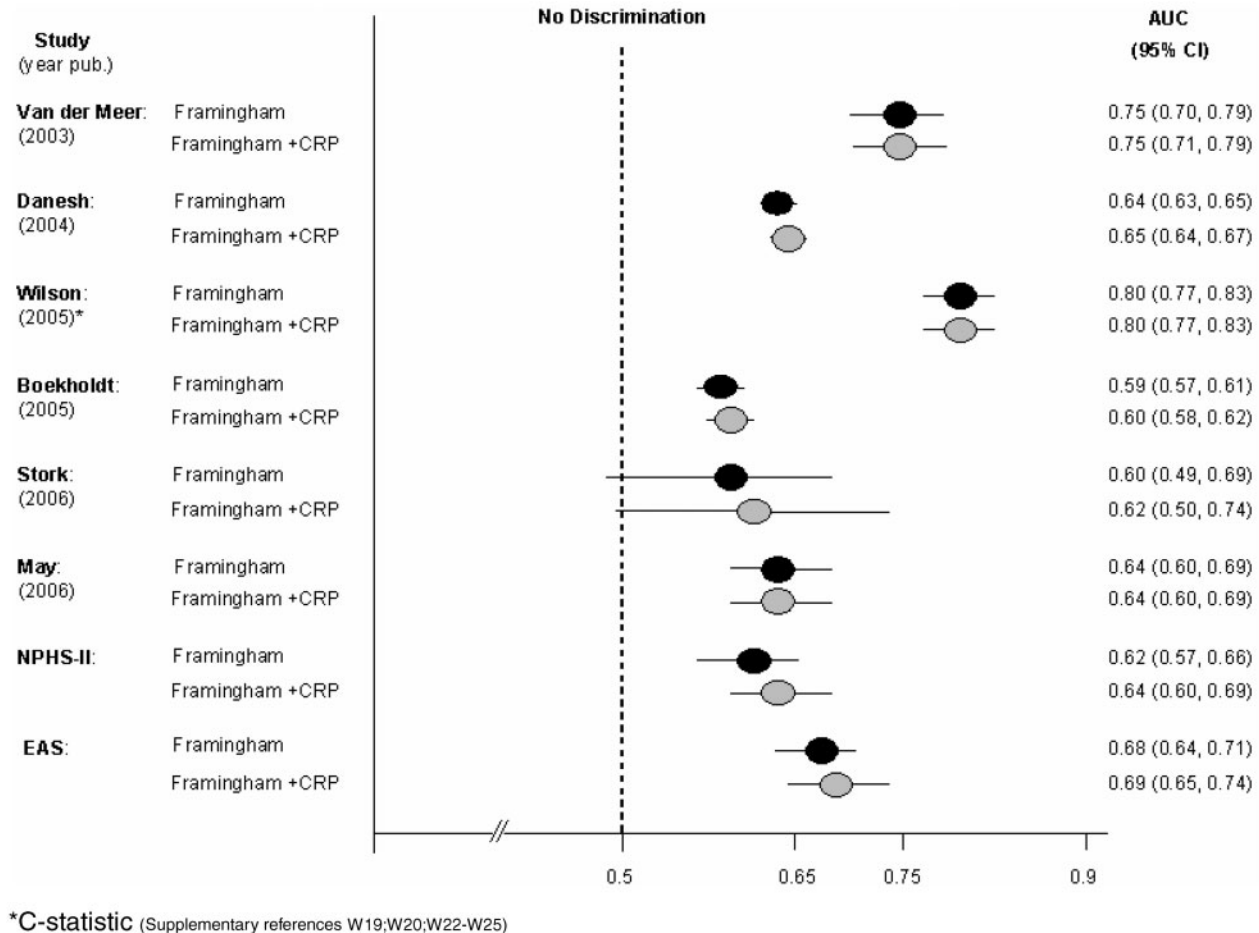
**Figure 3** Study-specific estimates of: (a) sensitivity (for a 10% false-positive rate; DR10) and (b) AUC for CRP obtained directly in NPHS-II, EAS and inferred from data extracted from 25 published reports. A sensitivity analysis by duration of follow up is also shown. The pooled values were calculated using the inverse variance-weighted method under random effect models

focused on event rates among individuals who shifted risk category on the addition of CRP, concluding that the actual risk of these individuals was more accurately assigned by the addition of CRP to the risk model. A comparison of observed risk in each category of predicted risk in NPHS-II and EAS (Table 4) that included all individuals in each category of predicted risk (i.e. those that shifted as well as those that did not) and a similar re-analysis of information from references W27 and 38 (data not shown) indicated that models that included or omitted CRP performed similarly. Moreover, the number of individuals that were reclassified between two risk categories based on a threshold 10-year CVD risk of 20% in WHS was 40 (0.15%) individuals who shifted above the threshold and 30 subjects (0.11%) who moved below the threshold. In the remaining 26 857, reclassification would not have affected decisions on the prescription of statins for primary prevention if these were based on a 20% CVD risk threshold. It was

not possible to assess the number of individuals for whom reclassification was appropriate using the NRI measure, as the data were not presented separately by case-control status.

## Discussion

In a new analysis of two prospective cohorts and a critical appraisal of published studies of CRP, we found consistent reporting of measures of association relevant to aetiological analysis, but variable reporting of measures more directly relevant to the predictive utility of a marker. Collation of the available evidence, drawing together information on the different measures of prediction (both reported and imputed) indicated that while consistently associated with CHD risk, CRP measurement provides variable, generally modest, though sometimes statistically significant improvement in risk prediction. This analysis helps rationalize the available evidence to help



**Figure 4** Observed AUC or C-statistic for Framingham-based models with and without the addition of CRP in NPHS-II, EAS and six published studies reporting these data with measures of dispersion. AUC or C-statistic for a further seven studies that did not report measures of dispersion were: Ridker 2002, 0.81 for Framingham alone and 0.81 for Framingham plus CRP; Koenig 2006, 0.735 for Framingham alone and 0.75 for Framingham plus CRP; St Pierre 2005, 0.74 for Framingham alone and 0.74 for Framingham plus CRP; Folsom 2006, 0.705 for Framingham alone and 0.706 for Framingham plus CRP; Cook 2006, 0.767 for Framingham alone and 0.77 for Framingham plus CRP; Sattar 2007, 0.813 for Framingham alone and 0.815 for Framingham plus CRP; Cao 2007, 0.63 for Framingham alone and 0.637 for Framingham plus CRP (Supplementary references W6, W8, W13, W15, W27, W28, W30)

clinicians, health providers and patients decide about whether or not measurement of CRP should be incorporated in the evaluation of CHD or CVD risk. The findings are also relevant to the conduct and reporting of studies of other emerging biomarkers being evaluated for CHD prediction and inform the debate on the optimal strategy for primary prevention of CHD.

## Predictive performance of CRP

### Discrimination

One recommendation of the 2006 NHLBI workshop on CRP was research to assess 'whether replacing a traditional risk factor(s) with CRP (or other newly discovered risk predictors) could improve CVD risk prediction', (<http://www.nhlbi.nih.gov/meetings/workshops/crp/report.htm>). In the NPHS-II and EAS

studies, measurement of CRP in individuals healthy at baseline provided limited discrimination for CHD events. CRP cut-points set to reduce misclassification (false-positive) rates to 5% or 10%, only detected between 10% and 20% of the people who eventually had events. The previously recommended CRP cut-point of 3 mg/l provided similarly modest discrimination. Discrimination over a range of cut-point values is summarized quantitatively as the AUC, which was 0.61 (0.57, 0.66) in NPHS-II, and 0.62 (0.57, 0.67) in EAS where 0.5 represents no discrimination and 1 represents perfect discrimination. Harrell's C-statistic was also not substantially different from the AUC in NPHS-II and EAS. The findings in NPHS-II and EAS were concordant with data obtained from prior reports identified by systematic review. Because of the variable reporting of the metrics relevant for assessing discrimination, these were derived, where



**Table 4** Observed event rate (risk) of CHD for different categories of predicted risk assessed using models incorporating or omitting CRP

|                | Category of predicted risk <sup>a</sup> |              |               |       | Proportion<br>Reclassified<br>(%) |
|----------------|---|--------------|---------------|-------|-----------------------------------|
|                | 0 to<br><5%                             | 5 to<br><10% | 10 to<br><15% | >15%  |                                   |
| <b>NPHS-II</b> |   |              |               |       |                                   |
| Without CRP    | 1119                                    | 1036         | 199           | 58    |                                   |
|                | 4.3%                                    | 6.9%         | 15.1%         | 17.2% |                                   |
| With CRP       | 1184                                    | 910          | 234           | 84    | 716/2412                          |
|                | 4.2%                                    | 6.6%         | 12.8%         | 22.6% | (29.7%)                           |
| <b>EAS</b>     |   |              |               |       |                                   |
| Without CRP    | 243                                     | 359          | 173           | 94    |                                   |
|                | 4.3%                                    | 7.8%         | 21.9%         | 22.8% |                                   |
| With CRP       | 333                                     | 331          | 146           | 109   | 254/919                           |
|                | 4.3%                                    | 8.4%         | 20.5%         | 26.1% | (27.6%)                           |

<sup>a</sup>Risk is 10-year risk in NPHS-II and EAS. Values in cells are total number of individuals in each predicted risk category, and 10 year event rate (%).

possible, using study-specific summary CRP values from cases and controls. The derived values for disease DRs and the AUCs for CRP (Table 3 and Figure 3) were concordant with those obtained by direct calculation in NPHS-II and EAS, providing evidence for consistency across studies.

Another proposed application is to add CRP to established predictors included in the Framingham model. In NPHS, EAS and in the 13 prospective studies that evaluated the effect of adding CRP to the Framingham risk equation (Figure 4), CRP provided either very small or no improvement in the AUC or C-statistic.

### Calibration

Calibration refers to the ability of a risk model to predict the observed risk in a group of individuals. Risk models are used clinically to target treatments for primary prevention to those at highest risk in order to enhance cost effectiveness and reduce unnecessary exposure to drugs. However, the addition of CRP to models based on established markers included in the Framingham equation did not substantially enhance the calibration in NPHS and EAS (Figure 1). Of the five published studies that evaluated the effect of adding CRP to the calibration of risk models, assessment was made in different ways but none appeared to identify a quantitatively large improvement in model fit, though statistically significant differences were sometimes judged as being clinically important.

### Reclassification

Because of concerns that neither discrimination nor calibration provides a fair test of a new marker, we

also studied the effect of adding CRP to established risk assessment methods using the NRI measure. Although this analysis breaks a continuous risk score into categories, it can be informative in assessing how the addition of a marker to a base model reassigns risk in those who do or do not suffer clinical events. When CRP was added to a Framingham-based model using four categories of 10-year CHD risk in NPHS-II and EAS, the proportion of subjects correctly reclassified (risk upgraded in eventual cases and risk downgraded in people remaining healthy) was almost matched by the proportion incorrectly reclassified, so that the NRI was only 8.5% (−1.3, 18.3) in NPHS-II and 8.8% (−1.3, 18.9) in EAS. When two risk categories were considered based on an individual traversing the 15% 10-year risk cut-point for initiating statin therapy for primary prevention, the NRI was reduced still further to 4.9% (0.8, 9.0) in NPHS-II and 3.0% (−3.0, 9.2) in EAS. Thus, in absolute terms, measurement of CRP in all 2479 men in NPHS-II would have led to a change in the decision to recommend statin therapy in only 11 individuals; statins would have been given to an additional 10 individuals, and withheld from one, and the cost effectiveness of such an approach has yet to be adequately assessed. Very similar results were observed in EAS (Table 2). Only one study previously reported the ability of CRP to reclassify subjects into risk categories. Only 70 of the 26 927 participants (0.26%) crossed the 10-year 20% CHD risk threshold following addition of CRP, but it was not possible from the data provided to derive the NRI in the absence of information on the eventual outcome among reclassified individuals. In NPHS-II and EAS, in direct comparisons of models that included or excluded CRP using reclassification tables that compare observed (actual) risk with predicted risk, there was very little overall difference in alignment between the two models (Table 4). Thus, overall, the available evidence from NPHS and EAS as well as published prospective studies indicated that the incremental predictive performance of CRP in CHD is limited when added to conventional risk factors, regardless of the metric utilized.

### Factors that constrain the predictive performance of CRP in CHD

Why should CRP, a sensitive marker of the inflammatory processes linked to atherosclerosis<sup>32</sup> and consistently associated with CHD risk, predict cases of CHD only modestly well? One important reason could lie in the fundamental relationships that determine the predictive performance of a marker. First, the distribution of CRP values in all populations studied was log-normal. Second, using pooled data from 18 studies, we identified a graded continuous relationship of CRP with risk of CHD, over the full range of CRP values observed in general populations, without a threshold. Thus, a substantial proportion of



all CHD events would be expected among the many individuals with near average levels of CRP, who are at intermediate risk of disease. This was observed in NPHS-II and EAS, as well as in 25 other studies that provided the relevant information. This likely contributes to the substantial overlap of CRP values among incident cases and controls (Supplementary Figure 2) that in turn, explains the small to moderate size of the association of CRP with CHD (an odds ratio of around two for a two SD difference in CRP). It also accounts for the difficulty in establishing a cut-point for CRP that adequately distinguishes those who suffered events, summarized quantitatively by the AUC.

Additional reasons are likely to contribute to the limited *incremental* utility for discrimination when CRP is added to standard risk assessment. First, it has been shown that combining markers that exhibit modest discrimination individually adds less than might be expected to the discrimination of a multi marker model.<sup>26</sup> Second, it is recognized that CRP is not only associated with blood pressure, LDL- and HDL-cholesterol, age and gender, but also with diabetes, smoking, left ventricular hypertrophy and atrial fibrillation, all of which already contribute to the Framingham risk model.<sup>33–35</sup> In some studies, the proportion of high CRP levels attributed to traditional risk factors was as high as a 78% in men and 67% in women.<sup>36</sup> This was borne out empirically in the 15 prospective studies that reported the AUC or the C-statistic for Framingham model with and without CRP (Figure 4). It is likely that similar constraints limit the degree to which addition of CRP recalibrates risk models or improves classification.

### Limitations and strengths

Some of the limitations of our study are worthy of consideration. First, not all studies reported the full range of metrics relevant for prediction. This was, in large part, the motivation for our critical appraisal. Nevertheless, we found no systematic differences in the characteristics of the studies reporting different metrics. Moreover, the fundamental relationships that constrain marker performance, namely the log-normal distributions of CRP values overall and the moderate, graded log-linear association with risk were more widely reported and were consistent across all studies. Where we derived metrics relevant to prediction from the published studies, we used aggregate rather than participant-level data with the assumption that the CRP distribution is log normal among both incident cases and controls. However, all the derived metrics and estimated relationships in published data were directly verifiable in EAS and NPHS-II, and all were closely concordant. Participant-level information would allow more detailed analyses, and this may become possible through resources being built by the Emerging Risk Factors Collaboration. We did not formally assess the possibility of publication bias in

this analysis, although this was addressed by a previous meta-analysis.<sup>6</sup> However, unpublished studies, should they exist, are more likely to attenuate, rather than inflate the estimates of predictive utility. Lastly, although CRP may have limited utility in the screening or prediction of CHD, there is also immense interest in the possibility that it may be causally involved in the development and progression of atheroma or its complications after an acute event<sup>37</sup> and our analysis was not designed to address this. Nor did our analysis address the utility of CRP for risk stratification around the time of an acute coronary syndrome, nor its potential pathological role in ischaemic tissue damage.

### Implications for evaluation of other novel markers

Twenty-five years ago more than 240 'risk factors' for CHD had already been identified, and many new markers have since been added. The pace of discovery of new markers is expected to accelerate with the new -OMICS technologies, and it is likely that one or more of these newly identified markers will be advocated for risk prediction in the same way as CRP. However, it is imperative that the utility of any marker is assessed using the appropriate metrics and not simply based on tests of 'independent' association that provide insufficient evidence on predictive utility. Better appreciation of the fundamental epidemiological relationships that constrain the performance of all novel markers will also be required; namely, the distribution in a population and the shape and strength of the association with CHD.

### Conclusions

In summary, new studies and a systematic review of published data indicate that while CRP is consistently associated with CHD and CVD risk, measurement of CRP provides much more limited information for risk prediction than tests of association alone might indicate. Previous guidance on the clinical use of CRP measurement may require updating in the light of these findings. Analyses of the type reported in this paper for CRP, could be also conducted for other novel biomarkers, to test whether or not any of these are likely to help better predict the occurrence of coronary events. All such analyses would benefit from the development of reporting standards to facilitate critical evaluation of the performance of biomarkers advocated for risk prediction.

### Supplementary data

Supplementary data are available at *IJE* online.

## Funding

UK Medical Research Council (to NPHS-II); the US National Institutes of Health (grant NHLBI33014); Du Pont Pharma, Wilmington, USA; British Heart Foundation (to EAS); the British Heart Foundation (PhD Studentship FS/02/086/14760 to T.S.); British Heart Foundation Programme Grant (PG2000/015 to J.A.C. and S.E.H.); Scottish Executive (Chief Scientist Office to I.T.); British Heart Foundation Shillingford Training Fellowship (FS/07/011 to R.S.); British Heart Foundation Senior Fellowship (FS/05/125 to A.D.H.).

## Acknowledgements

A.D.H. acknowledges generous support from the Rosetrees Trust.

**Conflict of interest:** None declared.

## References

- 1 Grundy SM, Cleeman JI, Daniels SR *et al.* Diagnosis and management of the metabolic syndrome: an american heart association/national heart, lung, and blood institute scientific statement. *Circulation* 2005;**112**:2735–52.
- 2 Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;**14**:32–38.
- 3 Vasan RSM. Biomarkers of cardiovascular disease: molecular basis and practical considerations. *Circulation* 2006;**113**:2262–335.
- 4 Danesh J, Collins R, Appleby P *et al.* Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: meta-analyses of prospective studies. *JAMA* 1998;**279**:1477–82.
- 5 Danesh J, Whincup P, Walker M *et al.* Low grade inflammation and coronary heart disease: prospective study and updated meta-analyses. *Br Med J* 2000;**321**:199–204.
- 6 Danesh J, Wheeler JG, Hirschfield GM *et al.* C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N Engl J Med* 2004;**350**:1387–97.
- 7 Pearson TA, Mensah GA, Alexander RW *et al.* Markers of inflammation and cardiovascular disease: application to clinical and public health practice: a statement for healthcare professionals from the centers for disease control and prevention and the american heart association. *Circulation* 2003;**107**:499–511.
- 8 Fortmann SP, Ford E, Criqui MH *et al.* CDC/AHA workshop on markers of inflammation and cardiovascular disease: application to clinical and public health practice: report from the population science discussion group. *Circulation* 2004;**110**:e554–59.
- 9 Myers GL, Rifai N, Tracy RP *et al.* CDC/AHA workshop on markers of inflammation and cardiovascular disease: application to clinical and public health practice: report from the laboratory science discussion group. *Circulation* 2004;**110**:e545–49.
- 10 Smith SC Jr, Anderson JL, Cannon RO III *et al.* CDC/AHA workshop on markers of inflammation and cardiovascular disease: application to clinical and public health practice: report from the clinical practice discussion group. *Circulation* 2004;**110**:e550–53.
- 11 Davey Smith G, Timpson N, Lawlor DA. C-reactive protein and cardiovascular disease risk: still an unknown quantity? *Ann Intern Med* 2006;**145**:70–72.
- 12 Wang TJ, Gona P, Larson MG *et al.* Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med* 2006;**355**:2631–39.
- 13 Musunuru K, Blumenthal RS. Biomarkers for prediction of cardiovascular events. *N Engl J Med* 2007;**356**:1472.
- 14 Ridker PM, Cook NR. Biomarkers for prediction of cardiovascular events. *N Engl J Med* 2007;**356**:1472.
- 15 Ware JH. Biomarkers for prediction of cardiovascular events. *N Engl J Med* 2007;**356**:1475.
- 16 Pepe MS, Janes H, Longton G *et al.* Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;**159**:882–90.
- 17 Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? *Br Med J* 1999;**319**:1562–65.
- 18 Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;**115**:928–35.
- 19 Pencina MJ, D'Agostino RB Sr, D'Agostino R Jr *et al.* Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–72.
- 20 Miller GJ, Bauer KA, Barzegar S *et al.* The effects of quality and timing of venepuncture on markers of blood coagulation in healthy middle-aged men. *Thromb Haemost* 1995;**73**:82–86.
- 21 Fowkes FG, Housley E, Cawood EHH *et al.* Edinburgh Artery Study: prevalence of asymptomatic and symptomatic peripheral arterial disease in the general population. *Int J Epidemiol* 1991;**20**:384–92.
- 22 Miller GJ, Bauer KA, Barzegar S *et al.* Increased activation of the haemostatic system in men at high risk of fatal coronary heart disease. *Thromb Haemost* 1996;**75**:767–71.
- 23 Waterworth DM, Talmud PJ, Bujac SR *et al.* Contribution of apolipoprotein C-III gene variants to determination of triglyceride levels and interaction with smoking in middle-aged men. *Arterioscler Thromb Vasc Biol* 2000;**20**:2663–69.
- 24 Younis J, Cooper JA, Miller GJ *et al.* Genetic variation in alcohol dehydrogenase 1C and the beneficial effect of alcohol intake on coronary heart disease risk in the Second Northwick Park Heart Study. *Atherosclerosis* 2005;**180**:225–32.
- 25 Tzoulaki I, Murray GD, Lee AJ *et al.* Inflammatory, haemostatic, and rheological markers for incident peripheral arterial disease: Edinburgh Artery Study. *Eur Heart J* 2007;**28**:354–62.
- 26 Wald NJ, Morris JK, Rish S. The efficacy of combining several risk factors as a screening test. *J Med Screen* 2005;**12**:197–201.
- 27 Law MR, Wald NJ, Morris JK. The performance of blood pressure and other cardiovascular risk factors as

- screening tests for ischaemic heart disease and stroke. *J Med Screen* 2004;**11**:3–7.
- <sup>28</sup> Pencina MJ, D'Agostino RB. Overall *C* as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;**23**:2109–23.
- <sup>29</sup> Wilson PWF, D'Agostino RB, Levy D *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;**97**:1837–47.
- <sup>30</sup> Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;**115**:92–106.
- <sup>31</sup> Cook NR, Buring JE, Ridker PM. The effect of including c-reactive protein in cardiovascular risk prediction models for women. *Ann Intern Med* 2006;**145**:21–29.
- <sup>32</sup> Ridker PM, Cushman M, Stampfer MJ *et al.* Plasma concentration of C-reactive protein and risk of developing peripheral vascular disease. *Circulation* 1998;**97**:425–28.
- <sup>33</sup> Pradhan AD, Manson JE, Rifai N *et al.* C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *JAMA* 2001;**286**:327–34.
- <sup>34</sup> Albert MA, Glynn RJ, Ridker PM. Plasma concentration of C-reactive protein and the calculated Framingham Coronary Heart Disease Risk Score. *Circulation* 2003;**108**:161–65.
- <sup>35</sup> Sesso HD, Buring JE, Rifai N *et al.* C-reactive protein and the risk of developing hypertension. *JAMA* 2003;**290**:2945–51.
- <sup>36</sup> Miller M, Zhan M, Havas S. High attributable risk of elevated C-reactive protein level to conventional coronary heart disease risk factors: the third national health and nutrition examination survey. *Arch Intern Med* 2005;**165**:2063–68.
- <sup>37</sup> Verma S, Szmítko PE, Ridker PM. C-reactive protein comes of age. *Nat Clin Pract Cardiovasc Med* 2005;**2**:29–36.
- <sup>38</sup> Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* 2007;**297**:611–19.

Published by Oxford University Press on behalf of the International Epidemiological Association

© The Author 2009; all rights reserved.

*International Journal of Epidemiology* 2009;**38**:231–234

doi:10.1093/ije/dyn353

# Commentary: C-reactive protein and risk prediction—moving beyond associations to assessing predictive utility and clinical usefulness

Ramachandran S Vasan<sup>1,2,3</sup>

Accepted 4 December 2008

*The only relevant test of the validity of a hypothesis is comparison of prediction with experience*

Milton Friedman

## Introduction

The last three decades have seen major advances in our understanding of the pathogenesis of coronary

heart disease (CHD), with substantial improvements in our management of patients at risk of disease and of those who have suffered a coronary event. The multifactorial nature of coronary disease and its evolution as a life course disease<sup>1</sup> due to the combined influences of genetic and lifestyle-related factors are firmly accepted concepts. We no longer argue about the role of cholesterol and its components as critical mediators of the disease, nor dispute whether the occlusive thrombus in the coronary tree in a patient with a myocardial infarction antedates or follows the clinical event. However, the debates in cardiology continue to rage with a shifting of the themes. One such contemporary controversy centres on if, when, and how novel

<sup>1</sup> The Framingham Study, Boston University School of Medicine, Framingham, MA, USA. E-mail: vasan@bu.edu.

<sup>2</sup> Department of Preventive Medicine, Boston University School of Medicine, Boston, MA, USA.

<sup>3</sup> Cardiology Section, Boston University School of Medicine, Boston, MA, USA.